

Diagrams, iconicity, and abductive discovery*

SAMI PAAVOLA

Abstract

In this article, the role of abductive reasoning within Peirce's diagrammatic reasoning is discussed. Both abduction and diagrammatic reasoning bring in elements of discovery but it is not clear if abduction should be a part of a fully developed diagrammatic system or not for Peirce. This relates to Peirce's way of interpreting abduction in his later writings. Iconicity and perceptual elements as a basis for discoveries are analyzed, both in deductive and abductive reasoning. At the end, the role of modern ideas of distributed cognition applied to the Peircean scheme is shortly delineated.

Keywords: diagrams; abduction; deduction; iconicity; discovery; distributed cognition

From the Peircean point of view, diagrams should be the heart of all reasoning. They are central in trying to understand the creative character of reasoning, especially because they are *iconic* signs. From early on, Peirce held that the iconic and observational character of all reasoning explained how deductive reasoning can draw necessary conclusions and, at the same time, present surprising discoveries (see below). According to Peirce, all necessary reasoning is diagrammatic (e.g., *EP* 2: 206, 212, 1903), and “diagrammatic reasoning is the only really fertile reasoning” (*CP* 4.571, 1906).

Diagrams, however, present a problem when it comes to Peirce's conception of the third mode of reasoning, i.e., *abduction*. Mathematical or necessary, that is, diagrammatic reasoning is the basis for all reasoning but does it include abduction? In reading Peirce's passages of this topic, it often seems that abduction (and, similarly, induction) is *not* a part of diagrammatic reasoning. And the same holds for modern commentators on Peirce's diagrammatic reasoning; usually they do not mention abduction at all or only in passing (Zeman 1964; Roberts 1973: 98–99, 127; Shin 1994; Allwein and Barwise 1996; Johnson-

Laird 2002). Is it even possible that abduction could be a part of diagrammatic systems (see Müller 1998: 92, 96)? Yet the exclusion from diagrammatic character is a bit strange if abduction is taken (as Peirce certainly did) as a third main mode of reasoning. And abduction is according to Peirce even more closely related to iconicity than deduction (*CP* 2.96, c. 1902, 1997: 276, 1903). Abduction is also supposed to be the way how “all the ideas of science come” (*EP* 2: 205, 1903). Is abduction then also a part of diagrammatic reasoning when it comes to surprising discoveries? How is “fertility” of reasoning related, on the one hand to diagrams, and on the other hand to abduction?

My aim in this paper is to suggest tentative solutions to these intricate questions. I am focusing on a Peircean claim that diagrammatic and iconic elements are a basis for creative aspects of human cognition in general, not so much concentrating on Peirce’s diagrammatic systems of logic as such (although I make some references to his existential graphs also; cf. Stjernfelt 2000; Hoffmann 2011). First, I briefly present the role of diagrams and iconicity as a basis for fertility within reasoning. Second, I explain some ways of understanding the relationship between abduction and diagrammatic reasoning. Third, I analyze the role of iconicity as a basis for abduction. Finally, I delineate the importance of distributed cognition for understanding abductive processes of discovery.

1. The nature of diagrammatic reasoning and iconicity

Before developing his system of existential graphs, Peirce already emphasized the *diagrammatic* character of all reasoning. Even simple syllogisms and formulas of algebra have their basis in constructing a diagram (an icon) and making observations and experiments upon this diagram. This iconic and observational character of deduction explains, according to Peirce, how deductive, *necessary* reasoning can also provide surprising discoveries.

It has long been a puzzle how it could be that, on the one hand, mathematics is purely deductive in its nature, and draws its conclusions apodictically, while on the other hand, it presents as rich and apparently unending a series of surprising discoveries as any observational science. Various have been the attempts to solve the paradox by breaking down one or other of these assertions, but without success. The truth, however, appears to be that all deductive reasoning, even simple syllogism, involves an element of observation; namely, deduction consists in constructing an icon or diagram the relations of whose parts shall present a complete analogy with those of the parts of the object of reasoning, of experimenting upon this image in the imagination, and of observing the result so as to discover unnoticed and hidden relations among the parts. (*EP* 1: 227, 1885)

Syllogisms and reasoning in general are about making diagrams and observations on the basis of these diagrams: “. . . in all reasoning there must be something amounting to a diagram before the mind’s eye, and that the act of inference consists in *observing* a relation between parts of that diagram that had not entered into the design of its construction” (*NEM* 4: 353, 893; *NEM* 4: 275–276, c. 1895)

Peirce’s basic definition of diagrams was that they are iconic signs concerning intelligible *relations* (in contrast to images which represent simple qualities and metaphors which concern representative character of signs; *CP* 2.277, c. 1902; see also *NEM* 4: 316, c. 1906). Diagrams do not necessarily resemble their objects in looks, but only in respect to the relations of their parts (*EP* 2: 13, 1895). Diagrams represent relations in between their own parts that are supposed to be analogous, i.e., similar to relations of parts within things represented (see *W* 5: 243, 1885).

This kind of a basic definition is a good starting point for understanding Peircean diagrams although some qualifications must be given also in relation to Peirce’s theory of signs. First, Peirce characterized icons at various times as if their fundamental characteristic is a (mere) *similarity*, or *resemblance* between the sign and its object (e.g., *EP* 1: 226, 1885; *EP* 1: 7, 1868; *EP* 2: 460–461, 1909). Similarity is a notoriously difficult concept in philosophy, and this has led to criticisms concerning this kind of an iconicity in many areas of research (see Stjernfelt 2000; Eco 2000). But similarity, upon deeper analysis, was not the most fundamental characteristic of icons for Peirce. Icons are signs that refer to their objects *merely with characters of their own even if there are no objects with similar characters*. So, in some sense, their starting point is just these signs in themselves, and similarity relations to objects is secondary.

An *Icon* is a sign which refers to the Object that it denotes merely by virtue of characters of its own and which it possesses, just the same, whether any such Object actually exists or not. It is true that unless there really is such an Object, the Icon does not act [as] a sign; but this has nothing to do with its character as a sign. Anything whatever, be it quality, existent individual, or law, is an icon of anything, in so far as it is like that thing and used as a sign of it. (*EP* 2: 291, 1903; see also *CP* 8.335, 1904; *EP* 2: 163, 1903; *CP* 4.447, c. 1903)

This “self-sufficient” character of icons and diagrams makes them fit well for reasoning.

A diagram, indeed, so far as it has a general signification, is not a pure icon; but in the middle part of our reasonings we forget that abstractness in great measure, and the diagram is for us the very thing. So in contemplating a painting, there is a moment when we lose the consciousness that it is not the thing, the distinction of the real and the copy

disappears, and it is for the moment a pure dream — not any particular existence, and yet not general. At that moment we are contemplating an *icon*. (EP 1: 226, 1885)

And they are a basis for new discoveries: “. . . a great distinguishing property of the icon is that by the direct observation of it other truths concerning its object can be discovered than those which suffice to determine its construction” (CP 2.279, c. 1895).

On the other hand, *pure* icons are fictitious, or ideal typical in nature; signs can actually be more or less iconic. Peirce maintained that a pure icon would be a possibility alone with just qualities (and Firstness by Peirce's categories) in contrast to “hypoicons,” which are *mainly* iconic (EP 2: 273, 1903; see also EP 2: 163, 1903). Also in relation to diagrams, although Peirce emphasized their iconic character, he also pointed out that they are of a mixed character, having also symbolic and indexical elements; a diagram is “. . . *predominantly* an icon of relations and is aided to be so by conventions. Indices are also more or less used” (CP 4.418, c. 1903, emphasis added; also CP 4.531, 1906; MS 293). Diagrammatic icons operate in connection with symbolic interpretants in logic (MS 293, c. 1906) So, it seems right to say that the interplay with several kinds of signs and with different kinds of elements within reasoning was central for Peirce, not just their iconic, indexical, or symbolic character (see EP 1: 226–228, 1885; EP 2: 10, 1894; EP 2: 441–442, 1908; cf. Barwise and Etchemendy 1996: 179–200).

2. The relationship between diagrams and abduction

Peirce emphasized that reasoning in general is diagrammatic, but does this mean all *deductive* reasoning, or induction and abduction as well? Abduction has characteristics closely connected to diagrammatic reasoning; *iconic* sign relationships are central in abduction, and abduction is closely connected to perceptual qualities, which explains why it is supposed to be an “originative” form of reasoning, that is, a way of introducing new ideas (EP 2: 226–241, 1903; Hanson 1958). But still it is not at all clear what, according to Peirce, is the relationship between abduction and diagrammatic reasoning.

Peirce never made any system on the basis of abduction. In his later writings, he presented abduction with a basic formula having a syllogistic form (see EP 2: 231, 1903):

The surprising fact, C, is observed;
But if A were true, C would be a matter of course.
Hence, there is reason to suspect that A is true.

On the other hand, abduction comes close to pure guessing (e.g., CP 7.219, 1901), or spontaneous conjectures (EP 2: 443, 1908). In his famous paper *Ne-*

glected Argument for the Reality of God (1908) he presented abduction (then named “retroduction”) as a form of “argument” rather than “argumentation” (*EP* 2: 441, 1908), meaning by argument “any process of thought reasonably tending to produce a definite belief” and by an argumentation “an [a]rgument proceeding upon definitely formulated premisses” (*EP* 2: 435, 1908). Within these formulations, abduction was aided by *il lume naturale*, i.e., a “light of reason” according to which human beings have a tendency to find true theories (*CP* 1.80–1.81, c. 1896). This tendency was not, Peirce thought, infallible but strong enough to help human beings to find fertile hypotheses better than by pure chance.

I think that these kinds of formulations by (later) Peirce made it more understandable why he did not include abduction with diagrammatic reasoning. Usually when Peirce described diagrammatic reasoning it was about *deductive reasoning*. He sometimes even *explicitly* wrote that induction and abduction are *not* diagrammatic reasoning — other than in an indirect way. Induction: “instead of experimenting on Diagrams . . . experiments upon the very Objects concerning which it reasons” (*MS* 293, c. 1906), whereas abduction is “. . . guessing . . . Such validity as this has consists in the generalization that no new truth is ever otherwise reached while some new truths are thus reached. This is a result of Induction; and therefore in a remote way Abduction rests upon diagrammatic reasoning” (*MS* 293, c. 1906).

Abduction, in the passages just cited, would then belong to pre-diagrammatic phases of the process of reasoning (see also Stjernfelt 2000: 372–373). Diagrams require “a perfectly consistent system of representation” (*CP* 4.418, c. 1903), whereas abduction for Peirce was more “open,” a way of suggesting good guesses.

On the other hand, there are some statements by Peirce that suggest he would have wanted to include abduction to diagrammatic reasoning:

. . . the limitation of the Graphs to the representation of necessary reasonings, which was imposed upon them in my “Prolegomena” [*Prolegomena to an Apology for Pragmaticism*, 1906], was needless: they are equally capable of representing the creations of explanatory conjectures, as well as the whole process of induction. (*MS* 296, c. 1907–1908; see also *MS* 296)

It seems then that Peirce’s aim of developing existential graphs further, and especially the gamma (or delta) part, meant that he wanted, at least in some of his formulations, to include abduction and induction in them. He did not manage to finish the gamma part, but it was supposed to include modalities that clearly presented a problem for the diagrammatic system according to Peirce (see *MS* 298, 1905). Modalities are important also from the point of view of abduction because with modalities he postulated a difference between deduction, induction, and abduction:

By *Deduction* one infers that if certain premisses are exactly true, then a certain conclusion must be true, either always or once in a certain proportion of cases in the long run. By *Adduction* [i.e., *Induction*] one infers that a certain state of things *is* true, at least approximately. By the third class of reasonings one only infers that a certain state of things *may be* true and that the indications of its being so are sufficient to warrant further examination. (*MS* 764; cf. *EP* 2: 216, 1903)

To successfully make graphs a “moving picture of thought” (see *MS* 298) seemed to require that abduction and induction should be included. But usually Peirce hesitated in drawing this conclusion (unless the inclusion is understood in an indirect way), stating that:

It may, however, be urged, that the Graphs only exhibit the processes of necessary ratiocination, but neither what passes in thought when from the contemplation of a bewildering mob of facts to an idea which might reduce them to order [i.e., abduction (SP)], nor yet that reverse operation by which, setting out from an idea, we make use of experiments to correct it and to confirm it [i.e., induction (SP)]. If this be made the ground of objection, the reply will be that although the rules which I have attached to the system of graphs only contemplate necessary reasoning, yet the other modes of inference are so related to the necessary mode that they are virtually given in the presentation of that. (*MS* 298, 1905)

In his later formulations of reasoning, abduction and induction were phases in a “methodetical” process (a methodological perspective), not just syllogistic forms of reasoning. With abduction a hypothesis is suggested concerning reality, which is made clear with deduction and tested with induction (*CP* 6.469–6.473, 1908; *CP* 7.202–7.219, 1901; see also, e.g., Paavola 2006b). Abduction and induction concern *real things* whereas deduction relates to ideal states of things: “. . . Abduction furnishes all our ideas concerning real things, beyond what are given in perception, but is mere conjecture, without probative force. Deduction is certain but relates only to ideal objects. Induction gives us the only approach to certainty concerning the real that we can have” (*CP* 8.209, c. 1905)

Abduction concerned mysteries surrounding the question of how human beings are and have been able to make conjectures about the world so successfully (*CP* 5.591, 1903; *EP* 2: 217, 1903).

Deduction proceeds by constructing a diagram representing purely ideal or hypothetical states of things, making observations about relations involved that suggest a hypothesis, and, on the basis of the hypothesis, making experiments (*CP* 2.778, 1901). There are stages in this process that seem to resemble abduction and induction, respectively; first, when on the basis of observations “a hypothesis suggests itself,” and, second, when the hypothesis is tested with experiments. But about the latter resemblance, Peirce comments explicitly:

“This is a proceeding extremely similar to induction, from which, however, it differs widely, in that it does not deal with a course of experience, but with whether or not a certain state of things can be imagined” (*CP* 2.778, 1901). So it seems that at least according to these kinds of formulations, within diagrammatic, deductive reasoning there are phases reminiscent of abduction and induction. Abduction and induction as such are *not* a part of diagrammatic reasoning but rather abduction precedes, and induction comes after diagrammatic reasoning. According to this line of thought, diagrams are involved in hypothetical and necessary mathematical universe, and abduction and induction are involved when things about our experience or reality are at the focus.

One way of interpreting abduction in connection to diagrammatic reasoning is then to see abduction from the point of view of deductive (necessary, diagrammatic) reasoning (Hoffmann 1999; see also Hintikka 1998). These elements close to abduction within deduction are connected especially to “theorematic reasoning.” Peirce made a distinction between corollarial and theorematic reasoning:

Corollarial deduction is where it is only necessary to imagine any case in which the premisses are true in order to perceive immediately that the conclusion holds in that case. All ordinary syllogisms and some deductions in the logic of relatives belong to this class. *Theorematic deduction* is deduction in which it is necessary to experiment in the imagination upon the image of the premiss in order from the result of such experiment to make corollarial deductions to the truth of the conclusion. (*NEM* 4: 38, 1902)

Theorematic reasoning is central for the “fertility” of deductive reasoning: “The peculiarity of theorematic reasoning is that it considers something not implied at all in the conceptions so far gained, which neither the definition of the object of research nor anything yet known about could of themselves suggest, although they give room for it” (*NEM* 4: 49, 1902)

On the basis of this connection, Michael Hoffmann has suggested that “[w]hat theorematic deduction is for mathematics, abduction seems to be for scientific discoveries in general” (1999: 292). Peirce himself did not usually make this connection between theorematic reasoning and abduction. Hoffmann has found just one place where there is a clear remark of this connection by Peirce:

But further study leads me to lop off a corollarial part from the Theorematic Deductions, which follows that part that originates a new point of view. This part of the theorematic procedure, I will call *theôric* reasoning. It is very plainly allied to retrodution [i.e., abduction (SP)], from which it only differs as far as I now see in being indisputable. (Peirce quoted in Hoffmann 1999: 293)

I agree with Hoffmann that there seem to be clear similarities between abduction and phases of theorematic reasoning. It is important, however, to note that even if there are some similarities, the forms of reasoning are still fundamentally different according to Peirce (theôric reasoning is “indisputable”). Deduction, even theorematic deduction, is about necessary, or apodictic reasoning, while abduction is supposed to be a weaker form of reasoning, coming close to guessing (Hoffmann 1999: 293–294). So, there is a need to understand the role of theorematic reasoning and abduction separately in relation to diagrammatic reasoning.

3. Abduction and iconicity

Peirce emphasized the iconicity of all reasoning (e.g., *NEM* 4: 353, 1893; *NEM* 4: 275, c. 1895; *NEM* 4: 158, 1903). The basis for this claim was that, according to Peirce, truths of reasoning are *perceived* in some important sense. Reasoning makes its conclusions *evident*, and this requires essentially iconic signs (*MS* 293, c. 1906). “It is . . . a very extraordinary feature of Diagrams that they *show* — as literally *show* as a Percept shows the Perceptual Judgment to be true — that a consequence does follow, and more marvelous yet, that it *would* follow under all varieties of circumstances accompanying the premises” (*MS* 293, c. 1906)

The observational character of reasoning is important also in bringing the element of *discovery* to reasoning (*EP* 1: 227, 1885).

The iconic character of reasoning is even more prominent in abduction. Abduction is inference “through an Icon” in contrast to induction and deduction, which are inferences, respectively, “through an Index” and “through a Symbol” (Peirce 1997: 276, 1903; see also *CP* 1.559, 1867). Abduction is closely related to perceptual judgments (*EP* 2: 226–241, 1903; Hanson 1958; Hoffmann 1999), it “shades into perceptual judgment without any sharp line of demarcation” (*EP* 2: 227, 1903).

It is, however, not clear how to interpret the relationship between iconicity, abduction, and perceptual judgments. In what sense is abduction supposed to be iconic and close to perception in contrast to indexicality of induction, and symbolicity of deduction (although it is not clear if induction is, according to Peirce, symbolic and deduction indexical — see Peirce 1997: 276–277, 1903; *EP* 2: 205–206, 1903)? I maintain that one central basis for this relationship is that abduction, iconic signs, and perceptual judgments all use *clue-like signs* as their starting point (Paavola 2005: 148). These clue-like characteristics mean that they are signs that only suggest a way (or ways) of interpreting them, or seeing a unifying connection between these clues. In abduction it means that a potentially promising way of arranging a previously confusing constellation

of things or characteristics is found or derived on the basis of these characteristics and background knowledge (Peirce 1997: 282–283, 1903; *EP* 2: 287, 1903).

At the first sight, perception and perceptual judgments seem to be a total opposite to this process. Percepts and also perceptual judgments *force* themselves on the perceiver; they are something that cannot be controlled by the person making them (*EP* 2: 155, 191, 1903). If something is perceived, it cannot be changed by the will. *But*, at the same time, perceptual judgments involve *interpretation*, which brings them close to abduction (*EP* 2: 299, 1903). This abductive or interpretative character of perception is evident in reversible figures (like the Necker Cube) where the same visual data can be interpreted in various ways (see *EP* 2: 228, 1903; see also Hanson 1958). Perception can be analyzed with a formula that is similar to abduction:

A well-recognized kind of object, M, has for its ordinary predicates P[1], P[2], P[3], etc., indistinctly recognized.

The suggesting object, S, has these same predicates, P[1], P[2], P[3], etc.

Hence, S is of the kind M. (*CP* 8.64, 1891)

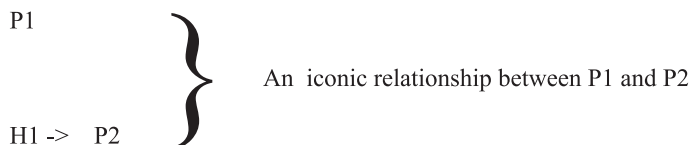
This sequence is abductive in form, but “[i]n perception, the conclusion has the peculiarity of not being abstractly thought, but actually seen, so that it is not exactly a judgment, though it is tantamount to one” (*CP* 8.65, 1891).

This brings us to *iconicity* in abduction. Abduction is an “originary” form of reasoning (*CP* 2.96, c. 1902), it should “cover all the operations by which theories and conceptions are engendered” (*CP* 5.590, 1903). Icons are central in abduction because icon is a sign that “denotes merely by virtue of characters of its own” (*EP* 2: 291, 1903 — see above). Icon is “an Original Sign . . . which is a Sign whose significant virtue is due simply to its Quality” (*CP* 2.92, 1902). In its pure form, icon is a mere possibility (*EP* 2: 277, 1903). “Icons . . . merely suggest the possibility of that which they represent, being, percepts *minus* the insistency and percussivity of percepts” (*MS* 293, c. 1906).

In short, icons, perceptual judgments, and abduction are different in important respects, because icons are, in their “pure” form, only pure possibilities, whereas perceptual judgments have their basis in the “insistency” of percepts, and abduction is a form of reasoning with consciously generated conclusions (at least in basic modes of reasoning). Still, there are important similarities between them. In all of them some characteristics or phenomena suggest a potential way of interpreting or explaining these characteristics or phenomena and bringing them into some kind of an order.

I think that a good starting point for understanding the relationship between abduction and iconicity is from 1901: “The mode of suggestion by which, in

abduction, the facts suggest the hypothesis is by resemblance — the resemblance of the facts to the consequences of the hypothesis.” (CP 7.218) I have depicted this in a simple diagram (figure 1):



∴ Maybe H1 (or something that is similar to H1)

Figure 1. *A basic formula of abduction with an iconic relationship*

So, according to this formula, an important iconic relationship within abductive inference is between some facts or phenomena stated (P1) and the consequences (P2) of a hypothesis suggested (H1). The abductive conclusion is that this kind of a hypothesis (H1) may be true (in relationship to those facts [P1]). A good example of this kind of abduction is provided by a case often referred in philosophy of science, that is, Ignaz Semmelweis' solution about the cause of the childbed fever (see more thoroughly Paavola 2006a). Semmelweis' crucial hypothesis (in his long search for the cause of this lethal decease) came when he noticed that the symptoms caused by blood poisoning are similar to symptoms of the childbed fever, and he made a hypothesis on the basis of this (which he then tested and modified); see figure 2:

P1 (symptoms of childbed fever)

H (blood poisoning) - > P2 (symptoms caused by blood poisoning)

P1 and P2 are similar (iconically)

∴ Maybe the cause of childbed fever is similar to blood poisoning (i.e., something detrimental was brought to mothers' bloodstream)

Figure 2. *An argument for Semmelweis' crucial hypothesis*

The abductive formula above (figure 1) could be presented differently so that it would emphasize the iconic character somewhat more¹ (figure 3):

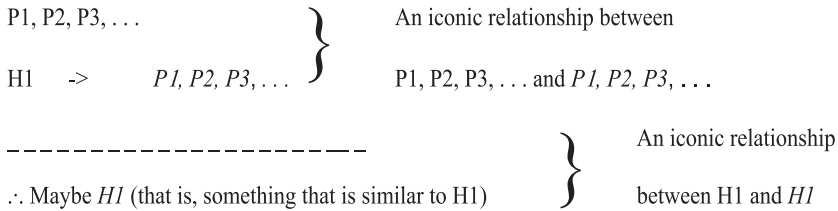


Figure 3. An abductive formula where iconic relationships are emphasized

So there is iconicity in two places here. I would like to emphasize a processual (or a “methodeutical,” in Peircean terms) approach as a way of analyzing the abductive search for hypotheses with clue-like signs. Semmelweis’ case can be taken as an example. A basic formula of abduction is often presented as if the surprising phenomenon is only a single fact (e.g., *EP* 2: 231, 1903) but a better picture is provided by emphasizing a constellation of facts or characteristics. Various features which were somehow problematic and without explanation were a starting point for the search for new hypotheses for Semmelweis; at the case of the crucial hypothesis symptoms of the childbed fever and the recognition that they are similar to symptoms in cases of the blood poisoning. This recognition was not a matter of course because these symptoms are quite complicated; lymphangitis, phlebitis, bilateral pleurisy, pericarditis, peritonitis, and meningitis, and they were not, of course, exactly alike with all patients. This is emphasized in the figure above by using different fonts (figure 3); Semmelweis could not count on conventional ways of interpreting the relationship between these two cases (that is, on “symbolic” similarity) because the conventional way of interpreting these symptoms was to see no connection between these two cases (the connection between childbed fever and blood poisoning was not recognized). By recognizing this similarity, Semmelweis used an iconic relationship, and concluded that there are good reasons to think that the cause *might be similar* in these two cases.

In order to understand the nature of abductive reasoning better, it is important to notice that the symptoms were not the only basis for Semmelweis to draw his conclusion. He knew many (other) curious and problematic things about childbed fever which guided his search for new hypotheses about its cause (he had tried many other explanations already before this “crucial” hypothesis). So I maintain that Semmelweis’ reasoning *can* be presented as above in the figure 2 but actually there were additional reasons for it which make it also understandable why Semmelweis noticed the connection between childbed fever and blood poisoning, while others did not. The reason was that Semmelweis had for long searched the cause of childbed fever, and this previous information constrained and guided his search. He, for example, knew that the two clinics of the hospital where he worked have a considerable degree of

difference in the mortality of the cases of childbed fever; that so-called street-births protected from the disease, and that there were time to time cases of the disease in rows. There was a sort of “proto-abduction” behind this search; Semmelweis knew various curious things related to childbed fever ($Q1, Q2, Q3, \dots$), he believed that if he could find a cause for childbed fever (X), it would probably explain at least most of these things, and his conclusion was that there possibly is to be found (there must be) such a cause (in relation to these curious things); figure 4:

$$Q1, Q2, Q3, \dots$$

$$X \rightarrow Q1, Q2, Q3, \dots$$

$$X?$$

Figure 4. A proto-abduction in search for X

And my point is *not* just that he used these curious things ($Q1, Q2, Q3, \dots$) as a preliminary test for his hypotheses (like for the crucial hypothesis above). He certainly did that also; after figuring out the hypothesis, he could reason that it seemed to give an explanation for these curious things: doctors who also did autopsies in a nearby mortuary caused most cases of childbed fever (without of course knowing it), and they operated in that clinic of the hospital where there were many cases of childbed fever, and for that reason there were time to time sickening in rows after doctors' examinations, and that was also why street-births were protected from childbed fever (they happened outside the hospital), etc. But what is important from an abductive point of view was that *before* testing, these curious things helped him to restrict in a preliminary (and abductive) way the search space for his hypotheses. He was, for example, searching for a hypothesis that would operate especially inside the hospital (births outside the hospital seemed to lessen the risk), and especially inside that clinic where doctors worked. These restrictions were tentative because there *were* cases of childbed fever also outside the hospital, and also in the other clinic.

I have used the notion of “similarity” as an essential basis for Semmelweis' abduction. Am I then contradicting myself because I tried to argue in section 1 that, for iconicity, similarity is *not* a basic notion but rather characters in themselves? Similarity must be somehow defined, and if it is defined then it seems to be a question of symbolicity rather than iconicity (see, e.g., Stjernfelt 1999, 2000: 358)?

I maintain, however, that both similarity and characters in themselves are important for abduction. Semmelweis had to take in account symptoms (i.e., characters) in themselves (both concerning symptoms of childbed fever and blood poisoning) in order to dissociate from the idea that childbed fever and blood poisoning would *not* have anything in common. If someone had suggested that symptoms of these two diseases have anything in common, most doctors at his time, and also Semmelweis before his recognition of this similarity, would have said that although there might be some superficial similarities between these two constellations of symptoms, they are totally different cases. While he had to take the characters of these symptoms as such (iconically), he had to see their similarity in some respects in order to make the abductive move. For abduction, it is enough to see a similarity in some respects in order to draw a hypothesis. And my argument above has been that for Semmelweis this hypothesis was promising because he saw that it matched to his earlier (tentative) restrictions and clues concerning the cause of childbed fever.

4. Diagrams and distributed abduction

In the previous section, I aimed at developing the ways of formulating abductive diagrams from the syllogistic point of view. In this section, I am briefly trying to give reasons for something that might look in a very opposite direction. I mean diagrams and abduction seen from the point of view of *distributed cognition*.

Distributed cognition refers to various approaches that have challenged the conception according to which individuals and mental constructions of individuals' minds are the center of intelligent activity (see, e.g., Salomon 1993; Hutchins 1995). As an alternative to standard problem-solving models, distributed cognition emphasizes social constructions developed in iterative processes by using external representations and features of the environment as an essential aid (Pea 1993: 65–67). For analytical purposes, physically, socially, and temporally distributed cognition can be distinguished although in actual inquiry processes all these aspects are connected together, and several modes and sources of knowledge are interacting with each other. *Physically distributed cognition* entails that the inquirers use external objects, features, artifacts, and tools to support their intelligent activity. *Socially distributed cognition* means those ways in which social interaction, collaboration, and social practices support activity. *Temporally* (and *culturally*) *distributed cognition* emphasizes the fact that human beings can use representations, tools, methods, etc., which are developed in long-term processes (also across generations) as an essential help.

Hoffmann (2011) has emphasized diagrammatic reasoning in general as a case of distributed cognition, which means that its function is “*to facilitate individual or social thinking processes in situations that are too complex as to being coped with exclusively by internal cognitive means*” (Hoffmann 2011). Diagrams are important not just because of their iconic character, but also because they are “objects” (or external representations) that can be perceived and observed (see *NEM* 4: 315, note 1, c. 1906) so that they can be processed not just by internal but also external cognitive means.

Peirce was a stern anti-psychologist in relation to logic (*MS* 498, 499), so things related to cognition should *not* have any consequences when it comes to abduction as a form of reasoning. With this Peircean spirit, my aim is not to claim that the ideas of distributed cognition have an influence on the *logic* of abduction but to maintain that Peirce’s concepts (especially when developed further) give means for conceptualizing processes of distributed cognition. Peirce’s basic definition of a sign can already be seen in relation to distributed cognition; signs refer to objects (cf. materially distributed cognition) and are interpreted in social processes (see Ransdell 2003; Bergman 2004: 40–80; cf. socially distributed cognition). Peirce emphasized, in general, social and “future-oriented” aspects of science and logic (e.g., *EP* 1: 52, 1868; *EP* 1: 81, 1869).

Along these lines, Peter Skagestad has developed Peirce’s “semeiotic model of mind” according to which the evolution of human beings proceeds especially by developing external artefacts. Skagestad has referred to Douglas C. Engelbart’s views on “augmentationism,” to Karl Popper’s view’s on evolutionary epistemology with external tools and artifacts but also to Peirce’s views on signs and knowledge (Skagestad 1993). Peirce sometimes emphasized the externality of signs and knowledge:

... psychologists undertake to locate various mental powers in the brain; and above all consider it as quite certain that the faculty of language resides in a certain lobe; but I believe it comes decidedly nearer the truth (though not really true) that language resides in the tongue. In my opinion it is much more true that the thoughts of a living writer are in any printed copy of his book than that they are in his brain. (*CP* 7.364, c. 1902; cf. also *CP* 2.54, c. 1902)

Although Peirce’s overall approach to inquiry and human cognition can then be interpreted through modern ideas about distributed cognition, his views concerning *abduction* emphasized inferential and/or instinctual aspects (Paavola 2005). Distributed cognition in relation to abduction is apparently a new interpretation (see Oatley 1996: 135–139; Magnani 2001, Paavola and Hakkarainen 2005; Hoffmann 2011). I think that there are good reasons for developing abduction like this towards “*il lume culturale*” (see Bonfantini and

Proni 1983: 134; Bonfantini 1988: 1253–1254; cf. Apel 1981: 170–171) or towards “il lume *materiale*” (Paavola 2007) rather than just as “il lume *naturale*” (i.e., a guessing instinct as a basis for abduction). One central motivation for Peirce to develop abduction was to explain how human beings have so successfully and quickly managed to find fertile hypotheses and true theories (*CP* 5.591, 1903; *EP* 2: 217, 1903). Peirce argued that it could not have happened by chance because that way would have taken too much time in relation to human history, and he ended up arguing for a guessing instinct (which is by no means infallible but still strong enough to have helped inquirers to find successful theories and ideas).

I maintain that a central basis for abduction is that human beings are good at using various kinds of constraints and clues to help them while searching for new ideas. It is beneficial to make an analytic difference between abductive *instinct* and abductive *inference* although Peirce himself merged these in his later writings (Paavola 2005). Besides instinctual or inferential abduction *distributed abduction* can be discerned which focuses on those ways things emphasized within distributed cognition constrain and guide the search for new ideas. Ideas do not turn up from scratch, or just within inquirer’s mind but from an interaction with material environment, previously (culturally) developed ideas, methods, and theories, and in collaboration with other people, and often by developing and modifying these ideas and theories in very long-term processes (cf. Hanson 1958: 72, 88). Even when someone (or a group of people) is developing something radically new, the elements and ingredients for this come from the interaction with the material, social, and cultural environment. Human beings create external representations which are then developed and modified further by others (see Magnani 2001).

5. Conclusion

I have delineated the role of diagrams and iconicity in relation to abductive processes of discovery. It is not entirely clear how Peirce meant abduction to be taken in relation to diagrammatic reasoning. What is more evident is that Peirce wanted to emphasize iconicity within abduction. I think that iconicity is a central element of abduction, and I have tried to clarify this connection. In his later years, Peirce himself was widening his system of logical graphs, and there are indications that they might have encompassed abduction if there had been more time for him to develop these graphs. On the other hand, he emphasized abduction as a guessing instinct and as a spontaneous conjecture that seemed to alienate him from making more specific systems of abduction.

Diagrams are also important because they are a bridge for understanding the distributed nature of human cognition. Distributed cognition is especially

important in abduction if it is interpreted as a weak form of inference with which human beings are in actual situations of puzzlement searching fertile, new hypotheses for explaining and interpreting real phenomena. There are then various kinds of things that constrain, guide, and help the abductive search for new ideas, and all of them should be taken into account if the aim is to understand processes of discovery.

Notes

- * Thank you for Hal White on help and sharp comments on this paper
1. I thank Petri Mäenpää for bringing this issue up.

References

- Allwein, Gerard & Jon Barwise (eds.). 1996. *Logical reasoning with diagrams*. New York: Oxford University Press.
- Apel, Karl-Otto. 1981. *Charles S. Peirce: From pragmatism to pragmaticism*. Amherst, MA: University of Massachusetts Press.
- Barwise, Jon & John Etchemendy. 1996. Heterogeneous logic. In Gerard Allwein & Jon Barwise (eds.), *Logical reasoning with diagrams*, 179–200. New York: Oxford University Press.
- Bergman, Mats. 2004. *Fields of signification: Explorations in Charles S. Peirce's theory of signs* (Philosophical Studies from the University of Helsinki 6). Vantaa: Dark.
- Bonfantini, Massimo A. 1988. Paper introducing the panel on abduction: Four questions on abduction. In Michael Herzfeld & Lucio Melazzo (eds.), *Semiotic theory and practice: Proceedings of the third international congress of the IASS Palermo*, 1245–1257. Berlin & New York: Mouton de Gruyter.
- Bonfantini, Massimo A. & Giampaolo Proni. 1983. To guess or not to guess? In Umberto Eco & Thomas A. Sebeok (eds.), *The sign of three: Dupin, Holmes, Peirce*, 119–134. Bloomington: Indiana University Press.
- Eco, Umberto. 2000. *Kant and the platypus: Essays on language and cognition*. New York: Harcourt Brace.
- Hanson, Norwood Russell. 1958. *Patterns of discovery*. Cambridge: University Press.
- Hintikka, Jaakko. 1998. What is abduction? The fundamental problem of contemporary epistemology. *Transactions of the Charles S. Peirce Society* 34(3). 503–533.
- Hoffmann, Michael. 1999. Problems with Peirce's concept of abduction. *Foundations of Science* 4. 271–305.
- Hoffmann, Michael. (2011). Cognitive conditions of diagrammatic reasoning. *Semiotica* 186(1/4).
- Hutchins, Edwin. 1995. *Cognition in the wild*. Cambridge, MA: MIT.
- Johnson-Laird, P. N. 2002. Peirce, logic diagrams, and the elementary operations of reasoning. *Thinking & Reasoning* 8. 69–95.
- Magnani, Lorenzo. 2001. *Abduction, reason, and science: Processes of discovery and explanation*. New York: Kluwer / Plenum.
- Müller, Ralf. 1998. Peirce's existential graphs: First inquiries towards a proper interpretation. In J. van Brakel & M. van Heerden (eds.), *C. S. Peirce categories to Constantinople: Proceedings of the international symposium on Peirce, Leuven 1997*, 83–100. Leuven: Leuven University Press.

- Oatley, Keith. 1996. Inference in narrative and science. In D. R. Olson & N. Torrance (eds.), *Modes of thought: Exploration in culture and cognition*, 123–140. Cambridge: Cambridge University Press.
- Paavola, Sami. 2005. Peircean abduction: Instinct, or inference? *Semiotica* 153(1/4). 131–154.
- Paavola, Sami. 2006a. Hansonian and Harmanian abduction as models of discovery. *International Studies in the Philosophy of Science* 20(1). 91–106.
- Paavola, Sami. 2006b. *On the origin of ideas: An abductivist approach to discovery* (Philosophical Studies from the University of Helsinki 15). Vantaa: Dark. <http://ethesis.helsinki.fi/julkaisut/hum/filos/vk/paavola/> (accessed 11 April 2011).
- Paavola, Sami. 2007. Abductive logic of discovery with distributed means. In Olga Pombo & Alexander Gerner (eds.), *Abduction and the process of scientific discovery*, 47–62. Lisbon: Centro de Filosofia das Ciências da Universidade de Lisboa.
- Paavola, Sami & Kai Hakkarainen. 2005. Three abductive solutions to the Meno Paradox — with instinct, inference, and distributed cognition. *Studies in Philosophy and Education* 24(3–4). 235–253.
- Pea, R. D. 1993. Practices of distributed intelligence and designs for education. In Gavriel Salomon (ed.), *Distributed cognitions: Psychological and educational considerations*, 47–87. Cambridge: Cambridge University Press.
- Peirce, C. S. 1931–1966. *The collected papers of Charles S. Peirce*, 8 vols., C. Hartshorne, P. Weiss & A. W. Burks (eds.). Cambridge: Harvard University Press. [Reference to Peirce's papers will be designated CP followed by volume and paragraph number.]
- Peirce, Charles Sanders. 1967. Manuscripts in the Houghton Library of Harvard University, as identified by Richard Robin, *Annotated catalogue of the Papers of Charles S. Peirce*. Amherst: University of Massachusetts Press. [Reference to Peirce's manuscripts will be designated MS.]
- Peirce, Charles Sanders. 1976. *The new elements of mathematics*, 4 vols., C. Eisele (ed.). Berlin & New York: Mouton de Gruyter; Atlantic Highlands, NJ: Humanities Press. [Reference to Peirce's New Elements will be designated NEM followed by volume and page number.]
- Peirce, Charles Sanders. 1982–. *Writings of Charles S. Peirce*, 5 vols., M. Fisch, E. Moore & C. Kloesel (eds.). Bloomington: Indiana University Press. [Reference to Peirce's writings will be designated W followed by volume and page number.]
- Peirce, Charles S. 1992. *Essential Peirce: Selected philosophical writings*, vol. 1 (1867–1893), N. Hauser & C. Kloesel (eds.). Bloomington: Indiana University Press. [Reference to vol. 1 of *Essential Peirce* will be designated EP 1.]
- Peirce, Charles Sanders. 1997. *Pragmatism as a principle and method of right thinking*. P. Turrisi (ed.). Albany: State University of New York Press.
- Peirce, Charles Sanders. 1998. *Essential Peirce: Selected philosophical writings*, vol. 2 (1893–1913). Peirce Edition Project (eds.). Bloomington: Indiana University Press. [Reference to vol. 2 of *Essential Peirce* will be designated EP 2.]
- Ransdell, Joseph. 2003. The relevance of Peircean semiotic to computational intelligence augmentation, *SEED* 3(3). 5–36.
- Roberts, Don D. 1973. *The existential graphs of Charles S. Peirce*. The Hague & Paris: Mouton.
- Salomon, Gavriel (ed.). 1993. *Distributed cognitions: Psychological and educational considerations*. Cambridge: Cambridge University Press.
- Shin, Sun-Joo. 1994. *The logical status of diagrams*. Cambridge, UK: Cambridge University Press.
- Skagstad, Peter. 1993. Thinking with machines: Intelligence augmentation, evolutionary epistemology, and semiotic. *Journal of Social and Evolutionary Systems* 16(2). 157–180.
- Stjernfelt, Frederik. 1999. How to learn more: An apology for a strong concept of iconicity. In Troels Degn Johansson, Martin Skov & Berit Brogaard (eds.), *Iconicity: A fundamental problem in semiotics*, 21–58. Aarhus: NSU Press.

- Stjernfelt, Frederik. 2000. Diagrams as centerpiece of a Peircean epistemology. *Transactions of the Charles S. Peirce Society* 36(3). 357–384.
- Zeman, Jay J. (1964). *The graphical logic of C. S. Peirce*. Chicago: University of Chicago dissertation. <http://www.clas.ufl.edu/users/jzeman/> (accessed 11 April 2011).

Sami Paavola (b. 1964) is a post-doctoral researcher at University of Helsinki <sami.paavola@helsinki.fi>. His research interests include abductive methodology, discovery, theories of learning and human cognition, and distributed cognition. His publications include “Abduction as a logic and methodology of discovery: The importance of strategies” (2004); and “Models of innovative knowledge communities and three metaphors of learning” (2004).

Copyright of Semiotica is the property of De Gruyter and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.